DOCUMENT RESUME

ED 061 021 RE 004 071

AUTHOR Felsenthal, Norman A.; Felsenthal, Helen
TITLE Utilizing the Computer to Assess the Readability of Language Samples.
PUB DATE Apr 72
NOTE 11p.; Paper presented at the meeting of the American Educational Research Association, Chicago, Ill., Apr. 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Automatic Indexing; Comparative Analysis; *Computational Linguistics; Information Processing; *Readability; Reading Level; *Reading Materials; *Scoring Formulas; *Word Frequency

ABSTRACT

A computer program called TEXAN (Textual Analysis of Language Samples) was developed for use in calculating frequency of characters, words, punctuation units, and stylistic variables. Its usefulness in determining readability levels was examined in an analysis of language samples from 20 elementary tradebooks used as supplementary reading materials. Three 200- to 300-word samples were selected to represent the beginning, middle, and end of each book. The TEXAN program was used to analyze the 60 samples according to four readability formulas: Gunning's "Fog" Index, Spache's Grade Level Indicator, Flesch's Reading Ease Index, and Flesch's Human Interest Index. Chi-square analysis and analysis of variance indicated that the samples were internally consistent. Relatively high correlations were found between the Gunning and Spache formulas, moderately low correlations were found between the Flesch formulas, and negative correlations were found between the two Flesch formulas and the Gunning and Spache formulas. It was concluded that the TEXAN program can be useful in analyzing readability, particularly when more than one formula is to be applied to a sample. Tables are included. (MS)

ED 060021

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

UTILIZING THE COMPUTER TO ASSESS THE READABILITY OF LANGUAGE SAMPLES[1]

Norman A. Felsenthal and Helen Felsenthal[2]
Purdue University

Readability formulas are valuable tools for the educator who wishes to assess the level or difficulty of various language samples. Utilizing the various formulas is a tedious and time-consuming task, however, as many who have classified and counted words and syllables will testify.

Research embodied in this paper demonstrates the employment of computer techniques to determine the readability levels of printed language samples. The research is also comparative in nature; i.e., four different readability indices were calculated for each sample and correlations between the indices were performed.

Methodology

TEXAN (Textual Analysis of Language Samples) is a computer program developed at Purdue University to calculate the frequency of characters, words, and punctuation units as well as other stylistic variables which are required to determine various readability formulas. The program has previously been used to determine the readability of radio commercial copy[3] and was employed in this study to analyze language samples from twenty elementary trade books (those intended for use in school libraries and for supplementary reading). Books selected for the research were from among the 1306 books cited by Eakin[4] and are listed alphabetically in Appendix B.

---

[1]Paper presented at the Annual Meeting of the American Education Research Association, Chicago, April 6, 1972.

[2]The authors are assistant professors in the Department of Communication and Education respectively.

[3]Norman Felsenthal, G. Wayne Shamo, and John R. Bittner, "A Comparison of Award-Winning Radio Commercials with Their Day-to-Day Counterparts," Journal of Broadcasting, XV (Summer, 1971), 309-315.

[4]Mary K. Eakin, Good Books for Children (Chicago: University of Chicago Press, 1962).

RE 004 071

More specifically, three 200 to 300-word samples were randomly selected from each of twenty books; i.e., each book yielded three language samples, one from the first third of the book, one from the middle portion, and a third from the final portion. The sixty language samples were key-punched and processed utilizing the textual analysis program.

Attributes of the Computer Program

Readability formulas are compounded from many different measurements. Some of these measurements are easily generated from computer algorithms; others can only be obtained through individual persual of the data. Those measurements readily quantified by the TEXAN program include statements, questions, exclamations, quotations, total words, total letters, letters per word, words per sentence, non-exempt words (those not included in any exempt list of up to 997 words), special words (those included in a specified list up to 100 words), large words (of "n" or more characters as specified), special endings (up to ten), words following a specified lead word (such as "the") and the frequency and percentage of individual characters (A, B, 1, 2, etc.).

One key element of many readability formulas which can not be directly obtained by the TEXAN program is syllable count. This can be accurately estimated, however, by dividing the number of letters in any message by 3.1127.[5]

---

[5]This ratio is derived from data analyzed in the study of radio commercial copy undertaken by the first author. In that study of forty language samples, correlations were run between a man-made syllable count and both letter count and vowel count. Characters per syllable were 3.1127 with a .98 correlation; vowels per syllable were 1.1761 with a .96 correlation. Coke and Rothkopf report a similar finding in which the Flesch Reading Ease Index was computed using man-made syllable counts and then re-computed using syllable counts based on estimates of vowels, consonants, and letters. Correlations between the Reading Ease scores utilizing the man-made syllable count and those using the computer-generated syllable count were .92 for vowels, .88 for letters, and .78 for consonants. See: Esther U. Coke and Ernst Z. Rothkopf, "Note on a Simple Algorithm for a Computer Produced Reading Ease Score," Journal of Applied Psychology, LIV (1970), 208-210.

Selection of the Readability Formulas

Between twenty and thirty readability formulas have been published and utilized over the past forty-odd years. From this multitude four formulas which utilized elements easily obtainable from the TEXAN program were selected for the analysis of the sixty language samples. These formulas are reproduced in Appendix A and include Gunning's "Fog" Index, Spache's Grade Level Indicator, Flesch's Reading Ease Index, and Flesch's Human Interest Index. The latter is not a readability measure in the strict definition of the term but does supply a level of reader interest which is helpful in evaluating language samples.

To generate the data necessary for these formulas, the TEXAN exempt word list was the Dale list[6] of 769 easy words (used in Spache) while the special word list consisted of the "personal words" designated by Flesch for his Human Interest Index.

All of the elements for these four formulas were generated by the TEXAN program with three exceptions. The Gunning Index requires a tabulation of words with three or more syllables. To accomplish this, the TEXAN program defined a "big word" as one containing eight or more letters. The "big word" list generated by TEXAN was then scrutinized to count those words with three or more syllables. Flesch's Reading Ease Index uses syllable count and these counts were estimated from letter count as previously described. The only measurement that required an examination of the raw data was the personal sentence count used in Flesch's Human Interest Index. Three of the five categories of personal sentences, quotations, questions, and exclamations, could be measured by TEXAN but the remaining two, commands and partial sentences, had to be tabulated by hand.

---

[6]This is the Clarence R. Stone revision of Edgar Dale's list of easy words as printed in: George Spache, Good Reading for Poor Readers, 3rd ed. (Champaign, Illinois: Garrard Publishing Co., 1962), 134-136.

Measurements were transferred to work sheets and a calculator was used to ascertain the readability for each passage. All 240 readability quotients (four formulas x three passages per book x twenty books) are reproduced in Appendix C. Mean scores for each book (the average of the three passages) are also included.

Analysis of the Data

The principal intent of this study was to demonstrate the feasibility of employing the computer to calculate readability formulas. Nevertheless, three distinct analyses were performed.

The first analysis examined the internal consistency of the readability in each of the twenty books selected for examination. Readability indices for the three samples from each book were compared and Chi-square procedures were employed to determine if readability variations within each book were excessive. None of the eighty Chi-squares (four indices x twenty books) were significant.[7]

Internal consistency of the twenty books as a single sample was measured in the second analysis. One-way analyses of variance were performed for each of four readability indices. None of the four indicated significant difference between the first, second, and third portions of the books. ANOVA data includes: Gunning ($F = .8122$, $df = 2, 59$); Spache ($F = .3902$, $df = 2, 59$); Flesch Reading Ease ($F = .1786$, $df = 2, 59$); and Flesch Human Interest ($F = .0464$, $df = 2, 59$).

---

[7]While other analysis supports the internal consistency of the twenty books, the non-significant Chi-squares could be attributed to a weakness in the experimental design. With only three samples for each book and, consequently, only two degrees of freedom in each Chi-square analysis, it was virtually impossible to obtain a significant Chi-square. In retrospect, the authors regret that they did not increase the numbers of samples from each book and, perhaps, decrease the number of books analyzed.

Scores from the four readability indices were correlated with one another for the third portion of the analysis. The correlation indicates a relatively high correlation

	Spache	Flesch Reading Ease	Flesch Human Interest
Gunning	.835	-.576	-.311
Spache		-.570	-.344
Flesch Reading Ease			.307

	Mean	Standard Deviation
Gunning	6.22	2.06
Spache	5.17	.79
Flesch Reading Ease	59.68	10.18
Flesch Human Interest	56.52	21.98

between the Gunning and Spache formulas, a moderately low correlation between the two Flesch formulas, and a negative correlation between the two Flesch formulas (higher index means greater reading ease/human interest) and the Gunning and Spache grade level indicators.

While correlation between Gunning and Spache was relatively high, a t-test of the means indicated a significant difference between the two measures ($t = 3.64$, $df = 59$). An examination of the data in Appendix C reveals that Gunning scores are frequently higher than Spache scores for those books in the upper elementary and junior high range while Spache scores are generally higher than Gunning scores for primary level books. This observation gives credence to the assertions by both authors concerning the respective grade levels for which the formulas are intended (1-3 for Spache, 6-12 for Gunning).

Conversely, while correlation between the two Flesch formulas was moderately low, a t-test of the means did not yield a significant difference ($t = 1.01$, $df = 59$).

Significance of the Research

Clearly, a computer program such as TEXAN can assist in the computation of readability formulas when a sizeable number of language samples must be analyzed. The computer program is particularily beneficial when two or more formulas are to be calculated from the same language sample. Not all readability formulas employ measurements easily obtained by computer but some do and others contain at least some elements that can be tabulated by computer.[8]

The trade books analyzed demonstrated an internal consistency which disproves the speculation stated by some educators that books become harder as they progress from beginning to end.

Readability formulas are interrelated but the relationship is not always a clear one. Considerable care must be taken to insure that the readability formulas chosen for analysis of certain language samples are appropriate for those samples.

---

[8]One formula which the authors have found particularly useful is Gillie's Abstraction Index. This quotient employs finite verbs, abstract nouns, and nouns preceded by the word "the." The last two measurements can both be made by TEXAN. See: Paul J. Gillie, "A Simplified Formula for Measuring Abstraction in Writing," Journal of Applied Psychology, XLI (1959), 214-217.

APPENDIX A

Readability Formulas Utilized in the Analysis

$$\text{Gunning}^{1} = \left(\text{words/sentence} + \frac{\text{words with 3 or more syllables}}{\text{total number of words}} \times 100\right) \times .4$$

$$\text{Spache}^{2} = .141 \times \text{words/sentence} + \left(.086 \times \frac{\text{non-extempt words}}{\text{total words}/100}\right) + .839$$

$$\text{Flesch Reading Ease}^{3} = 206.835 - \left(\frac{\text{syllables}}{\text{total words}} \times 100\right) - 1.015 \times \text{words/sentence}$$

$$\text{Flesch Human Interest}^{4} = \left(\frac{\text{personal words}}{\text{total words}} \times 363.5\right) + \left(\frac{\text{personal sentences}}{\text{total sentences}} \times 31.4\right)$$

---

[1]Robert Gunning, The Techniques of Clear Writing (New York: McGraw-Hill, 1952), 36-38. This formula, labeled a "Fog Index" by its author, produces a grade level indicator and is recommended for grade levels six through twelve.

[2]George Spache, "A New Readability Formula for Primary-Grade Reading Materials," Elementary School Journal, XLI (Fall, 1957), 214-217. The quotient is a grade level indicator and is recommended for the evaluation of primary books, grades one through three.

[3]Rudolf Flesch, How to Write, Speak and Think More Effectively (New York: Signet, 1963) 298-302. First published in 1948, the Flesch Reading Ease Index yields a score ranging from 0 to 100+ with higher scores indicating greater reading ease. No attempt is made to convert the scores to grade level; rather Flesch compares reading ease to seven levels of periodicals. Scores of 30 and below represent the most difficult level of reading typically found in scientific and professional journals. Scores of 90 to 100 represent the easiest level--comics. As mentioned in the body of this paper, syllable count was estimated on the basis of 3.1127 characters per syllable.

[4]Rudolf Flesch, How to Write, Speak and Think More Effectively (New York: Signet, 1963), 303-306. This quotient also yields scores ranging from 0 to 100+ with five intervals from dull to dramatic. Any score above 40 is labeled "highly interesting" and any score above 60 "dramatic." Personal sentences include questions, quotations, exclamations, commands, and partial sentences; personal words include personal pronouns, all nouns of gender, and the group words "people" and "folks."

APPENDIX B

Tradebooks Utilized in the Analysis

1. Agle, Nan H, and Wilson, Ellen. Three Boys and a Lighthouse. New York: Charles Scribner's Sons, 1951.
2. Anderson, C. W. Lonesome Little Colt. New York: MacMillan Co., 1961.
3. Ardizzone, Edward. Tim All Alone. New York: Henry Z. Walch, Inc., 1956.
4. Averill, Esther. Jenny's First Party. New York: Harper and Row, 1948.
5. Cavanna, Betty. Paintbox Summer. Philadelphia: Westminster, 1949.
6. Cleary, Beverly. Jean and Johnny. New York: Morrow, 1959.
7. _____. Otis Spofford. New York: Morrow, 1953.
8. George, Jean. My Side of the Mountain. New York: Dutton, 1959.
9. Henry, Marguerite. King of the Wind. Chicago: Rand McNally, 1948.
10. Jackson, Jacqueline. Julie's Secret Sloth. Boston: Little, Brown, 1953.
11. Lawson, Robert. The Great Wheel. New York: Viking, 1957.
12. Lenski, Lois. Cotton in My Sack. Philadelphia: Lippincott Co., 1949.
13. _____. Papa Small. New York: Henry Z. Walch, Inc., 1951.
14. _____. Peanuts for Little Ben. Philadelphia: Lippincott Co., 1952.
15. Rankin, Louise. Daughter of the Mountains. New York: Viking, 1948.
16. Sawyer, Ruth. Maggie Rose: Her Birthday Christmas. New York: Harper and Row, 1952.
17. Simpson, Dorothy. Island in the Bay. Philadelphia: Lippincott, 1956.
18. Tunis, John R. Highpockets. New York: Morrow, 1948.
19. Wilder, Laura I. Little House in the Big Woods. New York: Harper, 1953.
20. Wilson, Leon. This Boy Cody. New York: Franklin Watts, Inc., 1950.

APPENDIX C

Data from the Computer Analysis

Listed below in table format are the book numbers (see Appendix B for bibliographic data), the reported range (from Eakin, Good Books for Children), and the readability quotients for the three passages of each book. Mean scores are also reported.

Book #	Reported Range	Gunning	Spache	Flesch Reading Ease	Flesch Human Interest
1.	3-5	7.39	5.14	44.87	70.39
		4.54	4.80	61.33	86.91
		6.17	5.05	45.58	53.45
		6.03	4.99	50.59	70.25
2.	---	6.03	4.53	68.57	76.34
		5.20	4.43	17.00	48.06
		4.83	4.16	69.28	72.76
		5.35	4.37	51.61	65.72
3.	1-3	7.53	4.91	66.44	81.49
		7.15	4.94	72.62	65.35
		6.60	5.09	64.29	45.20
		7.09	4.98	67.78	64.01
4.	1-3	4.61	5.46	56.59	86.46
		5.10	4.84	57.69	37.01
		5.67	4.57	57.47	65.92
		5.12	4.95	57.25	63.13
5.	7-9	8.18	5.93	54.16	77.68
		7.75	4.85	53.69	77.55
		7.34	4.93	58.21	68.25
		7.75	5.23	55.35	74.49
6.	6-9	4.57	4.39	64.76	104.91
		5.49	4.82	61.17	71.74
		7.63	5.19	58.00	56.52
		5.89	4.80	61.31	77.72
7.	3-5	5.58	4.97	64.05	68.90
		5.07	4.49	70.01	95.10
		6.72	6.04	57.79	77.55
		5.79	5.16	63.95	80.51
8.	6-9	6.98	5.00	65.34	50.84
		5.95	5.06	67.43	51.24
		7.09	5.60	56.74	32.72
		6.67	5.22	63.17	44.93

9.	5-9	5.58	5.16	67.33	43.60
		5.82	5.08	66.82	42.61
		6.73	5.74	50.67	26.30
		6.04	5.32	61.60	37.50
10.	5-7	5.59	4.19	61.19	70.71
		4.35	5.09	60.39	71.99
		3.24	4.69	61.71	97.63
		4.39	4.65	61.09	80.11
11.	6-8	11.57	7.24	30.17	33.38
		9.55	7.16	45.87	21.08
		7.67	5.72	61.09	73.61
		9.59	6.70	45.71	42.69
12.	4-6	5.31	4.36	67.08	82.07
		4.71	5.02	64.60	69.71
		4.50	4.45	68.49	68.83
		4.84	4.61	66.72	73.53
13.	1-3	2.34	4.46	64.30	55.98
		3.40	4.91	61.58	51.62
		2.44	3.97	68.82	26.75
		2.72	4.44	64.90	44.78
14.	3-5	3.26	4.61	65.64	63.45
		2.74	4.10	71.58	106.29
		3.28	4.62	73.76	72.09
		3.09	2.90	70.32	80.61
15.	5-7	5.98	4.85	64.43	58.29
		9.92	6.19	48.50	14.54
		4.39	3.98	72.88	72.99
		6.76	5.00	61.93	48.60
16.	4-6	6.16	5.74	57.35	57.02
		6.05	5.09	65.23	66.37
		6.86	5.17	60.84	41.07
		6.35	5.33	61.14	54.82
17.	7-9	12.24	7.60	39.77	30.17
		5.32	4.82	65.35	70.00
		5.51	4.98	65.28	77.99
		7.69	5.80	56.80	59.38
18.	7-9	7.96	5.87	57.24	13.31
		7.50	5.18	59.70	26.61
		9.99	6.42	45.50	30.66
		8.48	5.82	54.14	23.52

19.	3-8	8.72	6.07	56.21	47.20
		5.82	5.44	69.21	71.38
		6.56	5.49	56.87	52.00
		7.03	5.66	60.76	56.86
20.	5-7	8.13	5.54	55.59	49.42
		6.10	5.30	64.01	66.48
		8.88	7.14	52.78	42.71
		7.70	5.99	57.46	52.87